

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

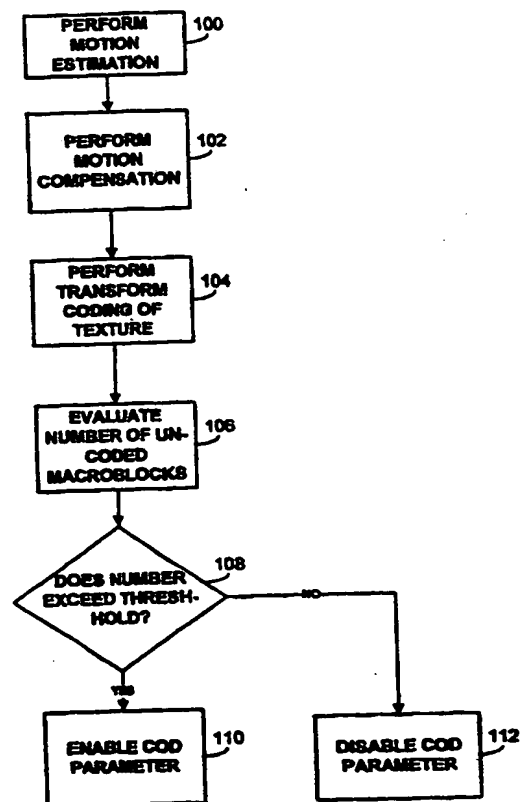
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 7/36, 7/50		A1	(11) International Publication Number: WO 99/34603
			(43) International Publication Date: 8 July 1999 (08.07.99)
(21) International Application Number: PCT/US98/20573			(81) Designated States: CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(22) International Filing Date: 30 September 1998 (30.09.98)			
(30) Priority Data: 09/001,573 31 December 1997 (31.12.97) US			
(71) Applicant: MICROSOFT CORPORATION [US/US]; One Microsoft Way, Redmond, WA 98052 (US).			
(72) Inventors: LEE, Ming-Chieh; 5558 166th Place, S.E., Bellevue, WA 98006 (US). CHEN, Wei-ge; 24635 S.E. 37th Street, Issaquah, WA 98029 (US).			
(74) Agent: MEYER, Joel, R.; Klarquist, Sparkman, Campbell, Leigh & Whinston, LLP, One World Trade Center, Suite 1600, 121 S.W. Salmon Street, Portland, OR 97204 (US).			

(54) Title: IMPROVED VIDEO CODING USING ADAPTIVE CODING OF BLOCK PARAMETERS FOR CODED/UNCODED BLOCKS

(57) Abstract

Video coding efficiency for high motion scenes is improved by adaptively disabling a parameter indicating whether texture and motion data is coded for a macroblock. The COD parameter is disabled when the number of macroblocks with substantially all zero motion and texture data is less than a threshold number. This reduces the number of bits required to code an interframe video image in a video sequence with high motion and large changes from frame to frame. The coded block pattern for chrominance is also used to determine how to perform entropy coding for the coded block pattern for luminance. In interframe blocks, if the chrominance blocks are coded, it is likely that the luminance blocks will be coded as well. The coded block pattern for chrominance, therefore, is used to select the appropriate entropy coding table for the coded block pattern for luminance.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IMPROVED VIDEO CODING USING ADAPTIVE CODING OF BLOCK PARAMETERS FOR CODED/UNCODED BLOCKS

5

FIELD OF THE INVENTION

The invention relates to video coding, and specifically, to an improved video coding method used in video coders and decoders.

BACKGROUND OF THE INVENTION

10 Full-motion video displays based upon analog video signals have long been available in the form of television. With recent advances in computer processing capabilities and affordability, full-motion video displays based upon digital video signals are becoming more widely available. Digital video systems can provide significant improvements over conventional analog video systems in creating, modifying, transmitting, storing, and playing full-motion video sequences.

15 Digital video displays include large numbers of image frames that are played or rendered successively at frequencies of between 30 and 75 Hz. Each image frame is a still image formed from an array of pixels based on the display resolution of a particular system. As examples, VHS-based systems have display resolutions of 320x480 pixels, NTSC-based systems have display resolutions of 720x486 pixels, 20 and high-definition television (HDTV) systems under development have display resolutions of 1360x1024 pixels.

The amounts of raw digital information included in video sequences are massive. Storage and transmission of these amounts of video information is infeasible with conventional personal computer equipment. Consider, for example, 25 a digitized form of a relatively low resolution VHS image format having a 320x480 pixel resolution. A full-length motion picture of two hours in duration at this resolution corresponds to 100 gigabytes of digital video information. By comparison, conventional compact optical disks have capacities of about 0.6 gigabytes, magnetic hard disks have capacities of 1-2 gigabytes, and compact optical 30 disks under development have capacities of up to 8 gigabytes.

To address the limitations in storing or transmitting such massive amounts of digital video information, various video compression standards or processes have

been established, including MPEG-1, MPEG-2, and H.26X. These video compression techniques utilize similarities between successive image frames, referred to as temporal or interframe correlation, to provide interframe compression in which motion data and error signals are used to encode changes between frames.

5 In addition, the conventional video compression techniques utilize similarities within image frames, referred to as spatial or intraframe correlation, to provide intraframe compression in which the image samples within an image frame are compressed. Intraframe compression is based upon conventional processes for compressing still images, such as discrete cosine transform (DCT) encoding. This type of coding is sometimes referred to as "texture" or "transform" coding. A
10 "texture" generally refers to a two-dimensional array of image sample values, such as an array of chrominance and luminance values or an array of alpha (opacity) values. The term "transform" in this context refers to how the image samples are transformed into spatial frequency components during the coding process. This use
15 of the term "transform" should be distinguished from a geometric transform used to estimate scene changes in some interframe compression methods.

Interframe compression typically utilizes motion estimation and compensation to encode scene changes between frames. Motion estimation is a process for estimating the motion of image samples (e.g., pixels) between frames.
20 Using motion estimation, the encoder attempts to match blocks of pixels in one frame with corresponding pixels in another frame. After the most similar block is found in a given search area, the change in position of the pixel locations of the corresponding pixels is approximated and represented as motion data, such as a motion vector. Motion compensation is a process for determining a predicted image and computing the error between the predicted image and the original image. Using
25 motion compensation, the encoder applies the motion data to an image and computes a predicted image. The difference between the predicted image and the input image is called the error signal. Since the error signal is just an array of values representing the difference between image sample values, it can be compressed using the same
30 texture coding method as used for intraframe coding of image samples.

Although differing in specific implementations, the MPEG-1, MPEG-2, and H.26X video compression standards are similar in a number of respects. The following description of the MPEG-2 video compression standard is generally applicable to the others.

5 MPEG-2 provides interframe compression and intraframe compression based upon square blocks or arrays of pixels in video images. A video image is divided into image sample blocks called macroblocks having dimensions of 16 x 16 pixels. In MPEG-2, a macroblock comprises four luminance blocks (each block is 8 x 8 samples of luminance (Y)) and two chrominance blocks (one 8 x 8 sample block
10 each for Cb and Cr).

In MPEG-2, interframe coding is performed on macroblocks. An MPEG-2 encoder performs motion estimation and compensation to compute motion vectors and block error signals. For each block M_N in an image frame N, a search is performed across the image of a next successive video frame N+1 or immediately
15 preceding image frame N-1 (i.e., bi-directionally) to identify the most similar respective blocks M_{N+1} or M_{N-1} . The location of the most similar block relative to the block M_N is encoded with a motion vector (DX,DY). The motion vector is then used to compute a block of predicted sample values. These predicted sample values are compared with block M_N to determine the block error signal. The error signal is
20 compressed using a texture coding method such as discrete cosine transform (DCT) encoding.

Object based video coding techniques have been proposed as an improvement to the conventional frame based coding standards. In object based coding, arbitrary shaped image features are separated from the frames in the video
25 sequence using a method called "segmentation." The video objects or "segments" are coded independently. Object based coding can improve the compression rate because it increases the interframe correlation between video objects in successive frames. It is also advantageous for variety of applications that require access to and tracking of objects in a video sequence.

30 In the object based video coding methods proposed for the MPEG-4 standard, the shape, motion and texture of video objects are coded independently.

The shape of an object is represented by a binary or alpha mask that defines the boundary of the arbitrary shaped object in a video frame. The motion of an object is similar to the motion data of MPEG-2, except that it applies to an arbitrary-shaped image of the object that has been segmented from a rectangular frame. Motion
5 estimation and compensation is performed on blocks of a "video object plane" rather than the entire frame. The video object plane is the name for the shaped image of an object in a single frame.

The texture of a video object is the image sample information in a video object plane that falls within the object's shape. Texture coding of an object's image
10 samples and error signals is performed using similar texture coding methods as in frame based coding. For example, a segmented image can be fitted into a bounding rectangle formed of macroblocks. The rectangular image formed by the bounding rectangle can be compressed just like a rectangular frame, except that transparent
15 macroblocks need not be coded. Partially transparent blocks are coded after filling in the portions of the block that fall outside the object's shape boundary with sample values in a technique called "padding."

Frame based coding techniques such as MPEG-2 and H26X and object based coding techniques proposed for MPEG-4 are similar in that they perform intraframe and interframe coding on macroblocks. The interframe coding format for these
20 techniques uses a special bit to indicate whether the interframe macroblock is coded. This special bit is sometimes called the COD bit or the "not coded" bit. To be consistent, we refer to this type of parameter as a COD bit or COD parameter. The COD bit indicates whether or not the encoded macroblock includes motion data and texture coded error data. In cases where the motion and error signal data is zero, the
25 COD bit reduces the information needed to code the macroblock because only a single bit is sent rather than additional bits indicating that the motion vector and texture data is not coded.

In addition to the COD bit, the coding syntax for macroblocks includes coded block parameters (CBP) indicating whether the coded transform coefficients
30 for chrominance and luminance are transmitted for the macroblock. If the transform coefficients are all zero for a block, then there is no need to send texture data for the

block. The Coded Block Parameters for chrominance (CBPC) are two bits indicating whether or not coded texture data is transmitted for each of the two chrominance blocks. The Coded Block Pattern for luminance (CBPY) are four bits indicating whether or not coded texture data is transmitted for each of the four
5 luminance blocks.

The CBPC bits are encoded along with another flag that provides information about the type of quantization for the macroblock. These flags are combined to form a parameter called MCBPC, and MCBPC is entropy coded using an entropy coding method such as Huffman or arithmetic coding. The CBPY flags
10 are also entropy coded using either Huffman or arithmetic coding.

While the COD bit has advantages in the coding of scenes with very little motion, it is inefficient for scenes that change frequently and have very few macro blocks with zero motion vectors (i.e. motion vectors indicating zero motion). Thus, there is a need for a more efficient application of the COD bit for these types of
15 scenes.

The variable length code for CBPY is based on the assumption that intraframe macroblocks include more coded luminance blocks than non-coded blocks, while for inter macroblocks, the opposite is true. This assumption is violated in some cases, and thus, leads to inefficient coding of the CBPY flags.

20

SUMMARY OF THE INVENTION

The invention provides an improved method of coding block parameters for video sequences that are coded using motion estimation and compensation. One aspect of the invention is a method for adaptively coding a COD parameter that
25 indicates whether texture and motion data is coded for interframe macroblocks. Another aspect of the invention is a method for adaptively selecting an entropy coding mode for the coded block parameters for luminance based on the coded block parameters for chrominance. Both of these features apply to object-based and frame-based video coding.

30 Adaptive coding of the COD parameter improves coding efficiency for a number of video coding applications, and in particular, for scenes with few

macroblocks with zero motion. Adaptive coding of the COD parameter refers to a method for selectively disabling COD parameter based on the number of non-coded macroblocks in a video image. Adaptive coding of the COD parameter is implemented in the encoder by determining the number of macroblocks in an image with zero texture and motion ("skipped macroblocks"). The COD parameter is then disabled for interframe macroblocks in the image when the number of skipped macroblocks is less than a threshold value.

The threshold number is derived from the total number of macroblocks in the image and the number of bits needed to encode motion and texture in an interframe macroblock in which the motion and texture data is all zero. At the threshold point, it takes fewer bits to code all macroblocks without the COD parameter because there are more COD bits than the number of bits needed to code motion and texture data for the interframe blocks for which the texture and motion data is zero. For example, if it requires five bits to encode zero motion and texture for a macroblock, then the threshold number is one-fifth the total number of macroblocks for the image.

The adaptive disabling of the COD parameter is represented using a COD disable flag that is provided with each interframe image. This flag indicates whether the COD parameter is included with each macroblock in the image. The encoder determines the value of this flag after performing motion estimation, motion compensation, and texture coding for the image because these processes indicate how many blocks have macroblocks with zero texture and motion data. The decoder uses the COD disable flag to determine whether to decode the COD parameter for each macroblock.

Coding efficiency can be improved further by adaptively selecting the entropy coding mode for the coded block pattern for luminance (CBPY) based on the coded block pattern for chrominance (CBPC). This improved coding method selects the entropy coding for CBPY in interframe macroblocks based on whether the CBPC bits are set for the macroblock. When the CBPC bits are set ("11") in an interframe macroblock, the luminance blocks are more likely to have non-zero transform coefficients. As such, entropy coding is selected so that fewer bits are

used to encode CBPY for coded luminance blocks. Conversely, when one of the CBPC bits is not set, entropy coding is selected so that fewer bits are used to encode CBPY for non-coded luminance blocks.

5 The methods summarized above improve coding efficiency for interframe macroblocks in video scenes with few macroblocks having zero motion and frequent changes from frame to frame. Adaptive disabling of the COD parameter can save up to 2-3% of the total bit rate for low bit rate applications. Adaptive selection of the CBPY coding mode for interframe blocks has been demonstrated to reduce the number of bits needed to encode CBPY by over 25%. These methods improve
10 coding efficiency without increasing the complexity of encoder and decoder design or creating compatibility problems.

Additional features and advantages of the invention will become more apparent from the following detailed description and accompany drawings of an implementation of the invention.

15

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of a video coder.

Fig. 2 is a block diagram of a video decoder.

Fig. 3 is a flow diagram illustrating an encoder process for adaptively
20 disabling the COD parameter for interframe macroblocks.

Fig. 4 is a flow diagram illustrating how a decoder interprets macroblocks for which the COD parameter is enabled and disabled.

Fig. 5 is a flow diagram illustrating how an encoder process manages entropy coding of CBPY parameters based on the values of CBPC parameters.

25 Fig. 6 is a flow diagram illustrating how a decoder interprets CBPY parameters of interframe macroblocks based on the values of the CBPC parameters.

DETAILED DESCRIPTION

Introduction

30 The first section below provides a description of a video encoder and decoder. Subsequent sections describe how to improve coding by: 1) adaptively enabling a block parameter indicating whether motion and texture is coded for a

block, and 2) improving the entropy coding of block parameters indicating whether luminance blocks are coded for a macroblock.

The invention applies to both frame-based and object-based video coding.

The invention improves the coding of a sequence of images, whether the images
5 represent video objects planes derived from a sequence of frames or the entire
frames. Object based coding includes similar motion and texture coding components
used in frame based coding as well as shape coding. The block syntax relevant to
the invention is similar in both frame based and object based coding. Thus, while
the encoder and decoder described in the next section are object based, they provide
10 a sufficient basis for explaining how to implement the invention in both frame based
and object based coding schemes.

Description of an Example Encoder and Decoder

Fig. 1 is a block diagram illustrating an implementation of an object-based
15 video encoder. The input 30 to the encoder includes a series of objects, their shape
information and bounding rectangles. The shape information, therefore, is available
before the encoder codes texture or motion data. Frame based coding differs in that
the entire frame is coded without shape information.

The shape coding module 32 receives the definition of an object including its
20 bounding rectangle and extends the bounding rectangle to integer multiples of
macroblocks. The shape information for an object comprises a mask or "alpha
plane." The shape coding module 32 reads this mask and compresses it, using for
example, a conventional chain coding method to encode the contour of the object.

Motion estimation module 34 reads an object including its bounding
25 rectangle and a previously reconstructed image 1636 and computes motion
estimation data used to predict the motion of an object from one frame to another.
The motion estimation module 1634 searches for the most similar macroblock in the
reconstructed image for each macroblock in the current image to compute the
motion data for each macroblock. The specific format of the motion data from the
30 motion estimation module 1634 can vary depending on the motion estimation
method used. For example, the motion data can include motion vectors or geometric

transform coefficients such as affine transform coefficients used to predict motion of the macroblock. In standard motion coding, the motion data includes one motion vector for each macroblock, or four motion vectors corresponding to the four luminance blocks in the macroblock.

5 The motion compensation module 38 reads the motion data computed by the motion estimation module and the previously reconstructed image 36 and computes a predicted image for the current frame. The encoder finds the difference between the image sample values in the input image block as specified in the input 30 and the corresponding sample values in the predicted image block as computed in the motion compensation module 38 to determine the error signal for the macroblock.

10 Texture coding module 40 compresses this error signal for inter-frame coded objects and compresses image sample values for the object from the input data stream 30 for intra-frame coded objects. The feedback path 42 from the texture coding module 40 represents the error signal. The encoder uses the error signal blocks along with the predicted image blocks from the motion compensation module to compute the previously reconstructed image 36.

15 The texture coding module 40 codes intra-frame and error signal data for an object using any of a variety of still image compression techniques. Example compression techniques include DCT, wavelet, as well as other conventional image compression methods.

20 The bitstream of the compressed video sequence includes the shape, motion and texture coded information from the shape coding, motion estimation, and texture coding modules. Multiplexer 44 combines and formats this data into the proper syntax and outputs it to the buffer 46.

25 While the encoder can be implemented in hardware or software, it is most likely implemented in software. In a software implementation, the modules in the encoder represent software routines executing in memory of a computer and memory used to store the video data. A software encoder can be stored and distributed on a variety of conventional computer readable media. In hardware implementations, the encoder modules are implemented in digital logic, preferably in an integrated circuit.

30 Some of the encoder functions can be optimized in special purpose digital logic

devices in a computer peripheral to off-load the processing burden from a host computer.

Fig. 2 is a block diagram illustrating a decoder for an object-based video coding method. A demultiplexer 60 receives a bitstream representing a compressed video sequence and separates shapes, motion and texture encoded data on an object by object basis. Shape decoding module 64 decodes the shape or contour for the current object being processed. To accomplish this, it employs a shape decoder that implements the inverse of the shape encoding method used in the encoder of Fig. 2. The resulting shape data is a mask, such as a binary alpha plane or gray scale alpha plane representing the shape of the object.

The motion decoding module 66 decodes the motion information in the bitstream. The decoded motion information includes motion data such as motion vectors for macroblocks blocks or geometric transform coefficients, depending on the type of estimation method used in the encoder. The motion decoding module 66 provides this motion information to the motion compensation module 68, and the motion compensation module 68 applies the motion data to previously reconstructed object data 70.

The texture decoding module 74 decodes error signals for inter-frame coded texture data and an array of color values for intra-frame texture data and passes this information to a module 72 for computing and accumulating the reconstructed image. For inter-frame coded objects, this module 72 applies the error signal data to the predicted image output from the motion compensation module to compute the reconstructed object for the current frame. For intra-frame coded objects the texture decoding module 74 decodes the image sample values for the object and places the reconstructed object in the reconstructed object module 72. Previously reconstructed objects are temporarily stored in object memory 70 and are used to construct the object for other frames.

Like the encoder, the decoder can be implemented in hardware, software or a combination of both. In software implementations, the modules in the decoder are software routines executing in memory of a computer and memory used to store the video data. A software decoder can be stored and distributed on a variety of

conventional computer readable media. In hardware implementations, the decoder modules are implemented in digital logic, preferably in an integrated circuit. Some of the decoder functions can be optimized in special purpose digital logic devices in a computer peripheral to off-load the processing burden from a host computer.

5

Improved Coding of COD Parameter

In current and proposed video coding standards, the parameter COD is used to indicate whether motion and texture are encoded for a interframe macroblock. The use of the COD bit is optimized for very low bitrate coding of low motion scenes. For many applications, and in particular for Internet applications, this use of the COD bit may be inefficient. For example, in these applications:

10

- The available bitrate for video is typically 10+kilobits per second (kbps);
- the picture size is typically QCIF (QCIF is a standard frame size that has 99 macroblocks);
- 15 • the frame rate is typically 5+ per second; and
- the scene change is frequent and motion between frames is mostly non-zero.

15

Under these conditions, the majority of the texture data for the macroblocks/blocks have non-zero coefficients. The COD bit is not effective in these circumstances because coding is skipped for few macroblocks yet the COD parameter is sent for all of the interframe macroblocks whether they are skipped or not. The bitrate for the COD parameter is about 0.5kbps. The method detailed below adaptively disables the COD bit and potentially saves 2-3% of the total bitrate.

20

To improve the efficiency of the COD parameter, the encoder adaptively disables this parameter for interframe coded images in which the number of non-coded macroblock exceeds a threshold value. In object based coding, the encoder adds a flag to the coding syntax at the video object plane level to indicate whether the COD parameter is disabled for all interframe macroblocks in the video object plane. The video object plane is an image representing an object in a frame.

25

In frame based coding, the encoder adds a flag to the coding syntax at the frame level to indicate whether the COD parameter is disabled for all macroblocks in

30

the frame. Note that in both the object based and frame based approach, the COD parameter is disabled for interframe macroblocks.

The video encoder of Fig. 1 encodes a video image in two primary stages. The first stage performs motion estimation and compensation, and the second stage constructs the bitstream in the appropriate coding syntax. Thus, the number of
5 macroblocks for which the motion vectors and texture data are all zero are determined in the first stage. Based on the number of these macroblocks, the second stage determines whether to disable the COD parameter for all of the macroblocks in the video image.

10 Fig. 3 is a flow diagram illustrating a process for adaptively disabling the COD parameter while encoding a video image. The encoder performs motion estimation and compensation (100, 102) and notes whether the motion vectors are zero. As it performs transform coding for each interframe macroblock and block (104), the encoder also notes whether the DCT transform coefficients for the
15 macroblock are all zero. This can occur in cases where the error values for all of the image samples in a macroblock are zero or are approximately zero.

When motion and texture coding is complete for the image, the second stage of the encoder determines the number of blocks for which the motion vectors and transform coefficients are all zero (106). These blocks are sometimes referred to as
20 skipped blocks because the encoding of motion vectors and transform coefficients is skipped for these macroblocks. If the number of skipped macroblocks exceeds a threshold value, then it is efficient to use the COD parameter for each macroblock, indicating whether the macroblock is coded. In this case, the COD parameter is enabled by clearing a COD disable flag for the image (110). Otherwise, the COD
25 disable flag is set for the image.

The coding syntax for the image includes the COD disable flag to indicate whether the macroblock syntax includes the COD parameter. The COD parameter is set to one if the transform coefficients for an inter macroblock are all zero, and otherwise, it is set to zero.

The following table will help illustrate how the COD disable flag improves coding efficiency for an image. The syntax for an interframe macroblock is as follows.

COD	MCBPC	CBPY	MVD	DCT block
-----	-------	------	-----	-----------

5 If the COD bit is set, then MCBPC, CBPY, MVD (motion vector data), and DCT block are not sent for the block. This is referred to as a skipped macroblock. Conversely, when the COD bit is not set in conventional coding standards, the macroblock includes entropy coded bits for MCBPC and CBPY, a coded motion vector or coded motion vectors (MVD), and at least one DCT block. Motion vectors
10 are typically coded as differential motion vectors. A differential motion vector is a vector quantity (e.g., has a vertical and horizontal component) representing the difference between two motion vectors, such as the difference between the motion vectors for the current and previous macroblock.

In scenes having few macroblocks with zero motion and frequent changes
15 from frame to frame, there are likely to be very few skipped macroblocks. Under these circumstances, the COD bit is inefficient because it adds an extra bit to the macroblock. If the COD parameter is disabled, each of the coded macroblocks have one less bit relative to the case where COD is enabled. Assuming, for example, that Huffman coding is used to encode MCBPC and CBPY, each of the macroblocks
20 with zero motion or texture data have extra bits for MCBPC (1 bit), CBPY (2 bits) and MVD (2 or more bits). The motion and texture for this type of macroblock requires at least 5 bits to code (MCBPC + CBPY + MV > 5). Thus, by disabling the COD parameter, this macroblock requires at least 5 bits to code, instead of just one bit in the case where COD is enabled.

25 The threshold number of skipped macroblocks can be evaluated analytically based on the total number of macroblocks in the image and the bits required to code a macroblock with zero motion and texture data and no COD parameter. The following expression gives an example of a threshold value:

30
$$(1/x) MB_{total} > MB_{skipped}$$

where MB_{skipped} is the number of skipped macroblocks, MB_{total} is the total number of macroblocks in the image, and x is the number of bits needed to code a macroblock when the motion and texture data is all zero. The threshold number is:

5 $(1/x) MB_{\text{total}}$.

It is important to emphasize that this threshold value is only an example. The number of bits needed to code a macroblock with zero motion and texture data can vary depending on the specific coding methods. For instance, in cases where Huffman coding is used, it is easier to estimate the number of bits needed to encode
10 MCBPC and CBPY as illustrated in the above example. In cases where arithmetic coding is used, it is more difficult to compute the number of bits needed to encode a macroblock with zero motion and texture data. However, a similar approach as shown above can be used in the case of arithmetic coding by using a conservative estimate of the number of bits needed to encode the parameters in a macroblock with
15 zero texture and motion data.

There are a variety of heuristic and analytical methods for selecting a threshold value. One way is to compute an average of the number of bits needed to encode macroblocks with zero motion and texture data for a frame and then use a formula similar to the one shown above (e.g., $(1/x_{\text{average}}) MB_{\text{total}}$) to compute a
20 threshold value for the frame. Another way is to compare, on a frame by frame basis, the number of bits needed to code the frame with and without the COD parameter and then select the approach that uses fewer bits.

The decoder has to be modified to support adaptive enabling and disabling of the COD parameter. Fig. 4 is a flow diagram illustrating how a decoder interprets
25 macroblocks for which the COD parameter is enabled and disabled. In a frame based approach, the COD disable flag is coded at the frame level, namely, there is a single flag for each frame. In an object based approach, the COD disable flag is preferably coded at the video object plane level. In other words, each of the interframe coded objects has a COD disable flag for each frame. When the decoder
30 decodes the frame or object level parameters, it decodes the COD disable flag (120) and determines whether the COD parameter is enabled or disabled (122).

If the COD parameter is enabled, the parameters at the macroblock level are interpreted as having a COD parameter for each macroblock in the image (124). In this case, the decoder selectively decodes motion vector and coded block parameters for each macroblock for which the COD parameter is set.

5 If the COD parameter is disabled, the parameters at the macroblock level are interpreted as not having a COD parameter for each macroblock (128). As such, the decoder decodes the motion vector and coded block parameter data for each macroblock.

Adaptive Coding of CBPY

10 The coding efficiency of scenes having few macroblocks with zero motion can be improved by adaptively selecting an entropy coding mode for the CBPY bits based on the CBPC bits. Fig. 5 is a flow diagram illustrating how an encoder process manages entropy coding of CBPY based on the values of CBPC.

15 The encoder evaluates the CBPC bits for each macroblock (130). There are two bits, one for each chrominance block. If both bits are set, meaning both chrominance blocks have non-zero transform coefficients, then it is likely that at least one of the four luminance blocks will have non-zero transform coefficients as well. The encoder adaptively selects an entropy code for CBPY based on the value of the CBPC parameters. If the CBPC bits are set, the encoder selects an entropy
20 encoding mode that assigns a higher probability to coded luminance blocks (134). Conversely, if at least one of the CBPC bits is not set, the encoder selects an entropy encoding mode that assigns a higher probability to un-coded luminance blocks (136). This approach is implemented by adaptively selecting between two entropy coding tables for inter-frame macroblocks based on the CBPC bits.

25 The conventional approach in proposed standards such as H263 is to use a first entropy coding table (e.g., a Huffman table) to create an entropy code for all intra-frame macroblocks, and a second entropy coding table to create an entropy code for all inter-frame macroblocks. The following table shows an example of entropy coding tables that are used to compute a variable length code (VLC). Table
30 1 is the conventional VLC table for intra-type macroblocks, and table 2 is the conventional VLC table for inter-type macroblocks. The CBPY bits indicate a one

(1) for a coded block, and zero (0) for an un-coded block. Note that un-coded blocks are deemed more likely for inter-macroblocks, and thus, are encoded with a code of shorter length than for coded blocks. Conversely, coded blocks are deemed more likely for intra-macroblocks, and thus, are encoded with a code of shorter length than for coded blocks.

5

Index	Table 1: CBPY (1 2 3 4)	Table 2: CBPY (1 2 3 4)	Number of Bits	Codes
0.00	0 0 0 0	1 1 1 1	4	0011
1	0 0 0 1	1 1 1 0	5	0010 1
2	0 0 1 0	1 1 0 1	5	0010 0
3	0 0 1 1	1 1 0 0	4	1001
4	0 1 0 0	1 0 1 1	5	0001 1
5	0 1 0 1	1 0 1 0	4	0111
6	0 1 1 0	1 0 0 1	6	0000 10
7	0 1 1 1	1 0 0 0	4	1011
8	1 0 0 0	0 1 1 1	5	0001 0
9	1 0 0 1	0 1 1 0	6	0000 11

Index	Table 1: CBPY (1 2 3 4)	Table 2: CBPY (1 2 3 4)	Number of Bits	Codes
10	1 0 1 0	0 1 0 1	4	0101
11	1 0 1 1	0 1 0 0	4	1010
12	1 1 0 0	0 0 1 1	4	0100
13	1 1 0 1	0 0 1 0	4	1000
14	1 1 1 0	0 0 0 1	4	0110
15	1 1 1 1	0 0 0 0	2	11

An improved method of coding CBPY is to select table 1 for inter-frame macroblocks when the CBPC blocks are both coded, and to select table 2 for inter-frame macroblocks when at least one of the CBPC blocks is not coded. Experiments have shown that adaptive selection of tables 1 and 2 for inter-frame macroblocks reduces the number of bits needed to encode CBPY for standard test video sequences from about a few percent to more than 26%. In the test case Akiyo where there is little change from frame to frame, adaptive selection of the CBPY tables produced a slight increase in the number of bits for CBPY (less than .1%). The advantages of this improved coding of CBPY are more significant for lower quantization step sizes and more complex scenes.

It is important to note that the above description of adaptive coding of CBPY is only one possible implementation. Note that there are three possible states for CBPC bits in the context of adaptive coding of CBPY: (1) both set (CBPC = 11); (2) both clear (CBPC = 00); and (3) one bit set-one bit clear (CBPC = 01 or 10). In

the implementation above, one entropy coding mode is used for state (1) and a second entropy coding mode is used for states (2) and (3). In some applications, it is advantageous to group state (2) with the entropy coding mode for state (1). A variety of different combinations of entropy coding modes and CBPC states are possible, and more than two entropy coding modes could also be used. For example, three different coding modes could be used for the three different states of CBPC.

The adaptive coding of COD and CBPY can be used together. Since the adaptive coding of CBPY is based on CBPC bits that are already in the macroblock format, it is not necessary to add an additional flag indicating whether this feature is enabled. However, adaptive coding of CBPY can be turned on and off by using a CBPY adaptive coding flag. One way to implement this flag would be to combine it with the COD disable flag. In this case, adaptive coding of COD and CBPY could be enabled and disabled together.

Adaptive coding of CBPY can be implemented in the decoder by selecting the table used to decode CBPY based on the CBPC bits. Fig. 6 is a flow diagram illustrating how a decoder interprets CBPY parameters of interframe macroblocks based on the values of the CBPC parameters. In a typical decoding operation, the decoder decodes the CBPC bits for a macroblock (140). It then determines whether $CBPC = 11$ (142). If so, the decoder decodes the VLC for CBPY using table 1. If not, the decoder decodes CBPY using table 2.

When implemented as described above, the improved coding of CBPY does not increase the complexity of the encoder or decoder and does not cause start code emulation concerns. "Start code emulation" refers to a compatibility problem caused by using a coding format that can be confused with a code used for synchronization called a "start code."

Conclusion

The coding methods described above improve coding efficiency for interframe blocks in two ways: 1) by adaptively disabling COD for scenes having few macroblocks with zero motion; and 2) by adaptively selecting entropy coding for CBPY using CBPC as indicator of motion and scene change. While the invention is described with reference to specific parameters, it is possible that the

format and definition of these parameters can change depending on the implementation. The specific definitions used above are useful because they are consistent with current and proposed video coding standards. However, it is also possible to use the invention in video coding methods that use a different syntax.

- 5 In view of the many possible implementations of the invention, it should be recognized that the implementation described above is only examples of the invention and should not be taken as a limitation on the scope of the invention. Rather, the scope of the invention is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

We claim:

1. In a video coder for coding texture and motion data of video images in a block format, a method for improving compression of the video images comprising:
 - 5 determining a value indicating how many blocks have motion and texture data values that are all substantially zero;
comparing the value to a threshold value;
when the value exceeds the threshold, using a parameter in the block format for blocks in a video image to indicate whether a corresponding block in the video
10 image includes texture and motion data; and
when the value does not exceed the threshold, encoding the block format for the blocks in the video image without using the parameter.
2. The method of claim 1 wherein the video images represent video objects
15 in an video sequence that are encoded using shape, texture and motion data.
3. The method of claim 1 wherein the video images represent video frames in video sequence that are encoded using texture and motion data.
- 20 4. The method of claim 1 wherein the parameter is a COD bit,
wherein the block format includes a COD bit for each block when the value exceeds the threshold, and the block format does not include a COD bit for any block in the video image when the value does not exceed the threshold.
- 25 5. The method of claim 4 further including:
wherein the video image is a video object in an object based coding method;
and
adaptively disabling or enabling the COD bit on an object by object basis in each frame.

6. The method of claim 4 further including:

wherein the video image is a video frame in a frame based coding method;

and

5 adaptively disabling or enabling the COD bit on a frame by frame basis such that the COD bit is either enabled for an entire frame or disabled for the entire frame.

7. The method of claim 1 further including:

10 encoding a disable bit with the video image to indicate whether the parameter is disabled.

8. A computer readable medium having instructions for performing the steps of claim 1.

15 9. In a video decoder, a method for decoding an inter-frame block comprising:

determining whether a coded macroblock parameter is enabled for a video image;

20 when the coded macroblock parameter is enabled, determining whether the coded macroblock parameter is set, and when the coded macroblock parameter is not set, skipping motion and texture decoding for a corresponding macroblock;

when the coded macroblock parameter is disabled, decoding motion and texture data for each macroblock in the video image.

25 10. The method of claim 9 wherein the video image is a frame in a video sequence and the coded macroblock parameter is enabled or disabled for each frame in the video sequence by coding a COD disable flag with each frame.

30 11. The method of claim 9 wherein the video image is a video object plane in a frame in a video sequence and the coded macroblock parameter is enabled or

disabled for each video object plane in the video sequence by coding a COD disable flag with each video object plane.

5 12. A computer readable medium having instructions for performing the steps of claim 9.

13. In a video coder for coding texture and motion data of video images in a block format, a method for improving compression of the video images comprising:
evaluating coded block parameters for chrominance for an interframe
10 macroblock;

adaptively selecting an entropy coding mode for the interframe macroblock including:

analyzing the coded block parameters for chrominance to estimate whether luminance blocks are likely to be coded for the macroblock;

15 employing a first entropy coding mode to encode the coded block parameters for luminance when the analysis of the coded block parameters for chrominance indicates that the luminance blocks are likely to be coded, wherein the first entropy coding mode uses fewer bits to indicate that the luminance blocks are coded than the number of bits used to indicate that the
20 luminance blocks are not coded; and

employing a second entropy coding mode to encode the coded block parameters for luminance when the analysis of the coded block parameters for chrominance indicates that the luminance blocks are not likely to be coded, wherein the second entropy coding mode uses more bits to indicate
25 that the luminance blocks are coded than the number of bits used to indicate that the luminance blocks are not coded.

14. The method of claim 13 including:

30 while encoding inter-Macroblocks for a video image, evaluating the coded block parameters for chrominance, and employing a first entropy code table to encode the coded block parameters for luminance when the coded block parameters

for chrominance are set, and employing a second entropy code table to encode the coded block parameters for luminance when at least one of the coded block parameters for chrominance is not set;

5 wherein the first entropy code table uses fewer bits to indicate that the luminance blocks are coded than the number of bits used to indicate that the luminance blocks are not coded; and wherein the second entropy code table uses more bits to indicate that the luminance blocks are coded than the number of bits used to indicate that the luminance blocks are not coded.

10 15. A computer readable medium having instructions for performing the steps of claim 13.

16. In a video decoder for decoding texture and motion data of video images in a block format, a method for decoding the video images comprising:

15 evaluating coded block parameters for chrominance for an interframe macroblock;

 selecting an entropy decoding mode for the interframe macroblock including:

 when the coded block parameters for chrominance indicate that the chrominance is coded for the interframe macroblock, employing a first
20 entropy decoding mode to decode the coded block parameters for luminance, wherein the first entropy decoding mode uses fewer bits to indicate that the luminance blocks are coded than the number of bits used to indicate that the luminance blocks are not coded; and

 when the coded block parameters for chrominance indicate that the chrominance is not coded for at least one chrominance block associated with
25 the interframe macroblock, employing a second entropy decoding mode to encode the coded block parameters for luminance, wherein the second entropy decoding mode uses more bits to indicate that the luminance blocks are coded than the number of bits used to indicate that the luminance blocks
30 are not coded.

17. The method of claim 16 wherein the first entropy decoding mode employs a first entropy coding table to decode the coded block parameters for luminance for the interframe macroblock, and the second entropy decoding mode employs a second entropy coding table to decode the coded block parameters for luminance for the interframe macroblock.

18. The method of claim 17 wherein the first and second tables are employed adaptively for each interframe macroblock based on the coded block parameters for chrominance in each interframe macroblock, such that the table used to decode each interframe macroblock can differ based on characteristics of each macroblock.

19. A computer readable medium having instructions for performing the steps of claim 16.

20. A computer readable medium having a data structure storing compressed interframe macroblocks; the computer readable medium comprising:
coded interframe images including coded interframe macroblocks corresponding to each of the interframe images;
each interframe image include a COD disable flag indicating whether the interframe macroblocks for the interframe image each have a COD parameter; wherein during decoding of the interframe images, the COD disable flag is operable to enable and disable decoding of the COD parameter such that when the COD parameter is enabled, texture and motion decoding for interframe macroblocks with the COD parameter set is skipped, and when the COD parameter is disabled, texture and motion decoding is performed for all interframe macroblocks in the interframe image.

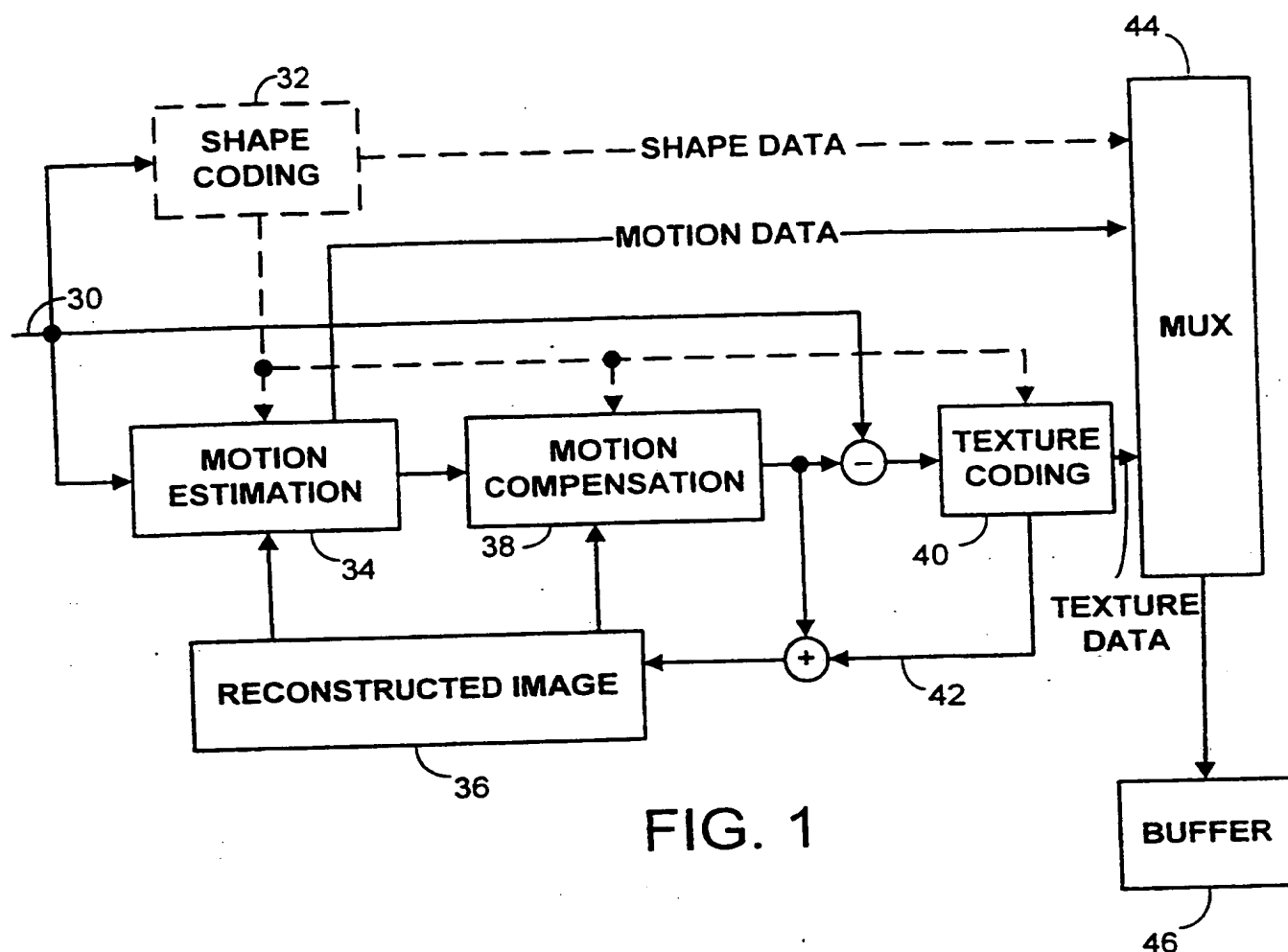


FIG. 1

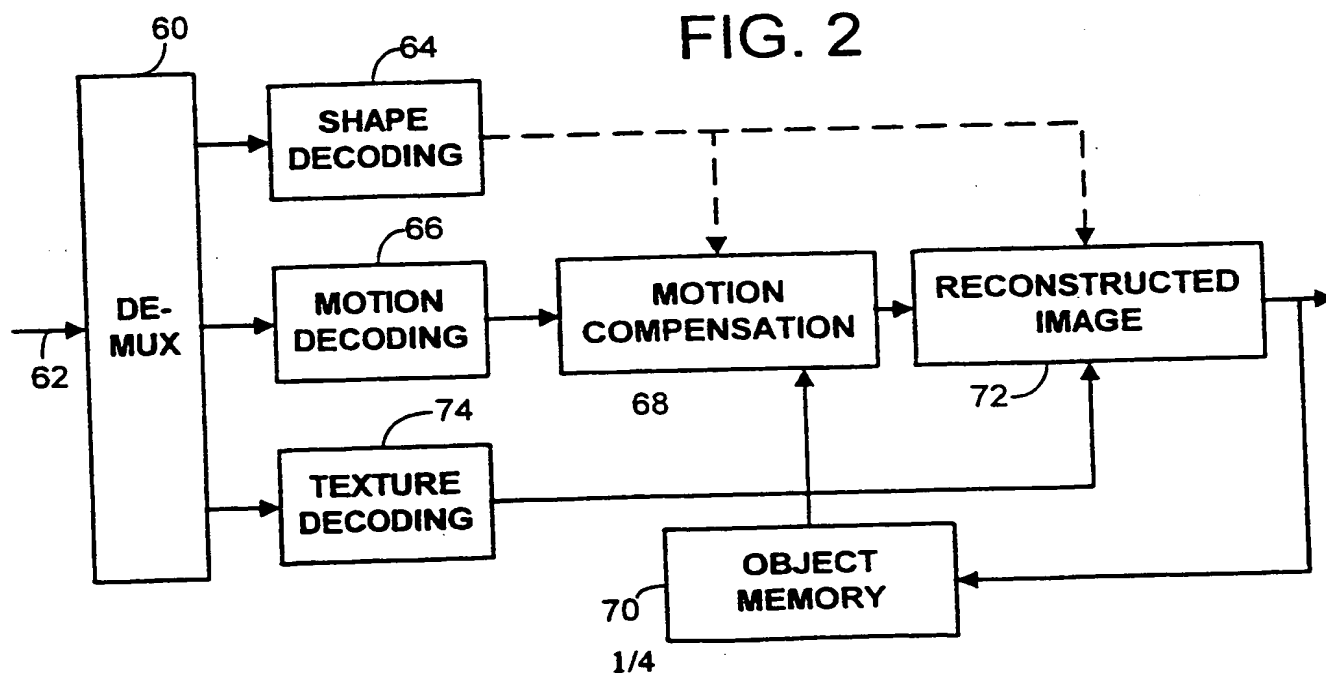


FIG. 2

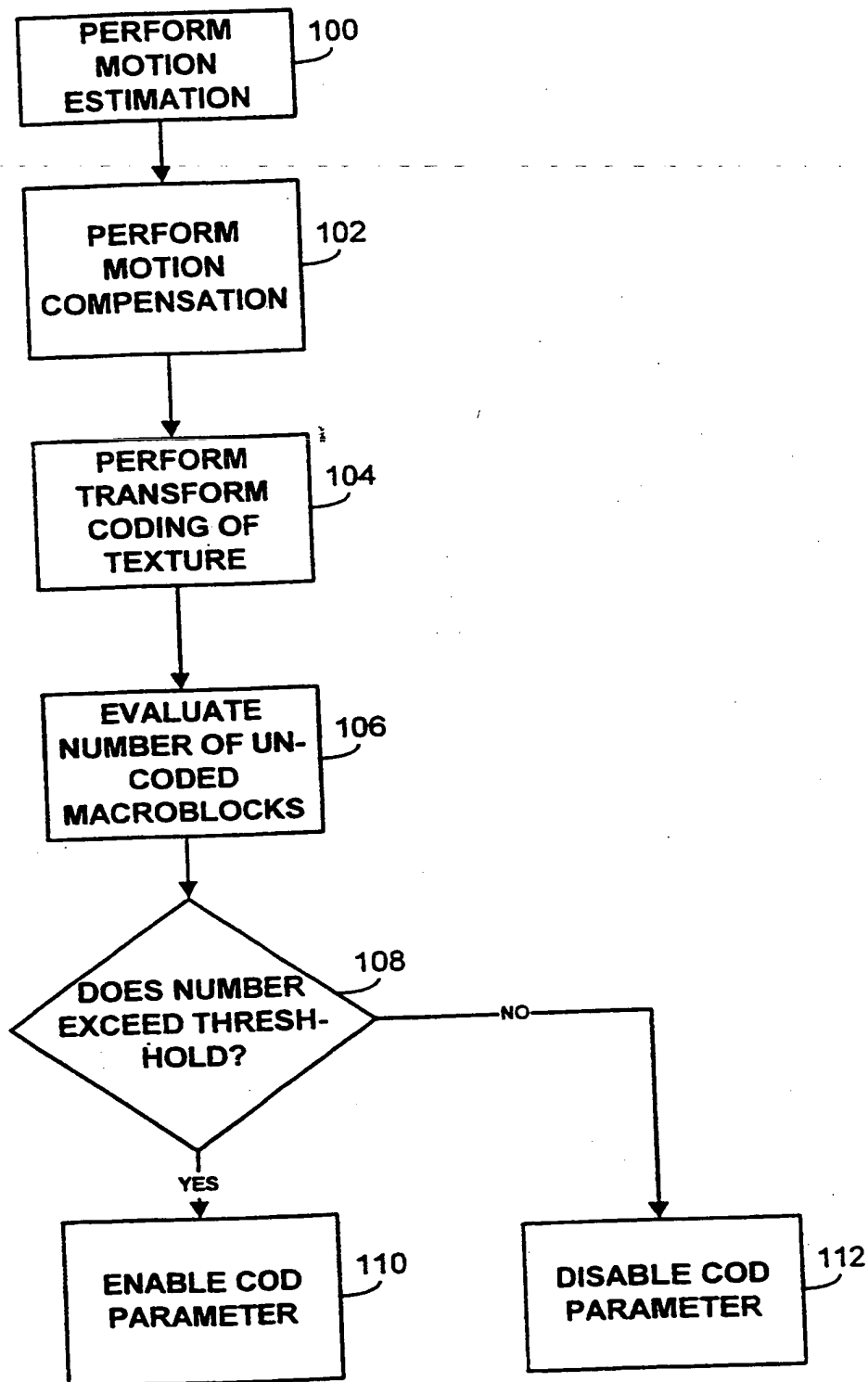


FIG. 3

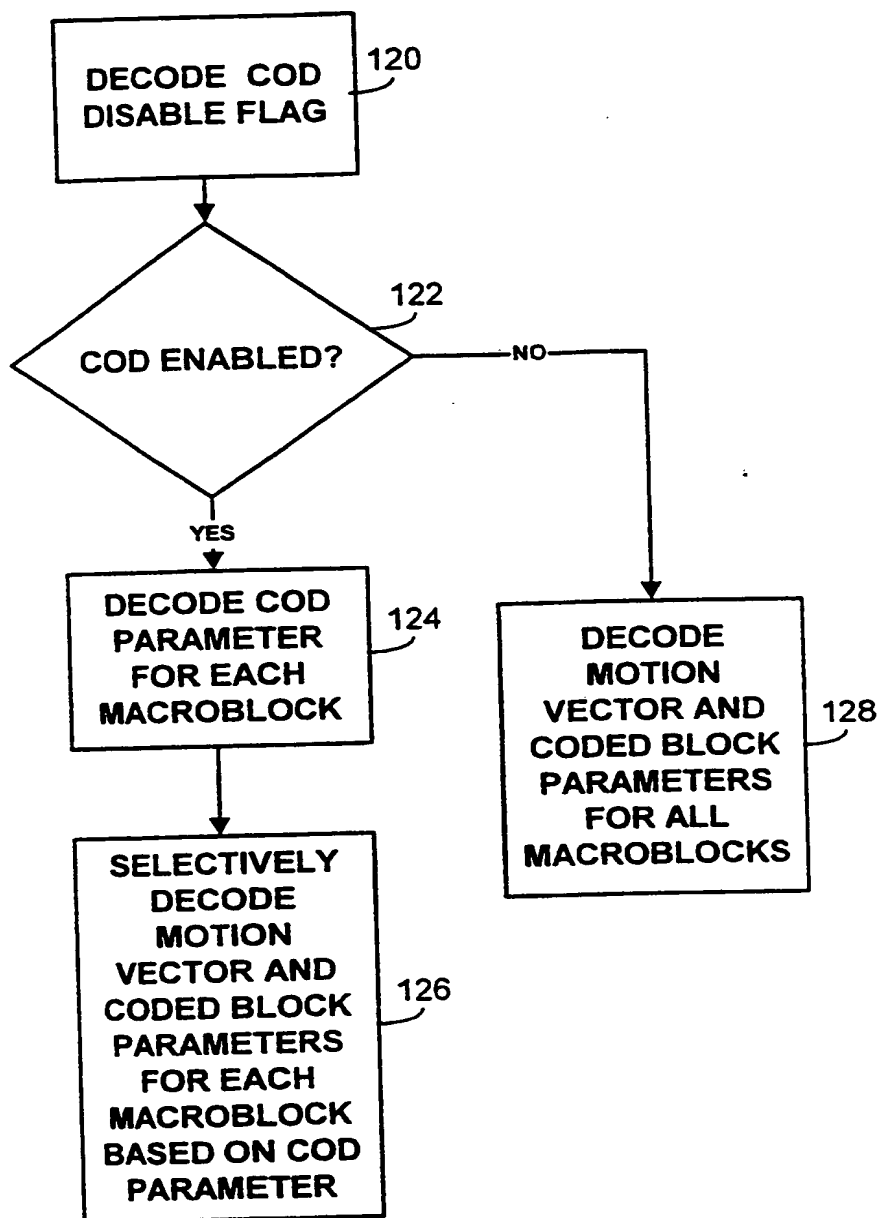


FIG. 4

FIG. 5

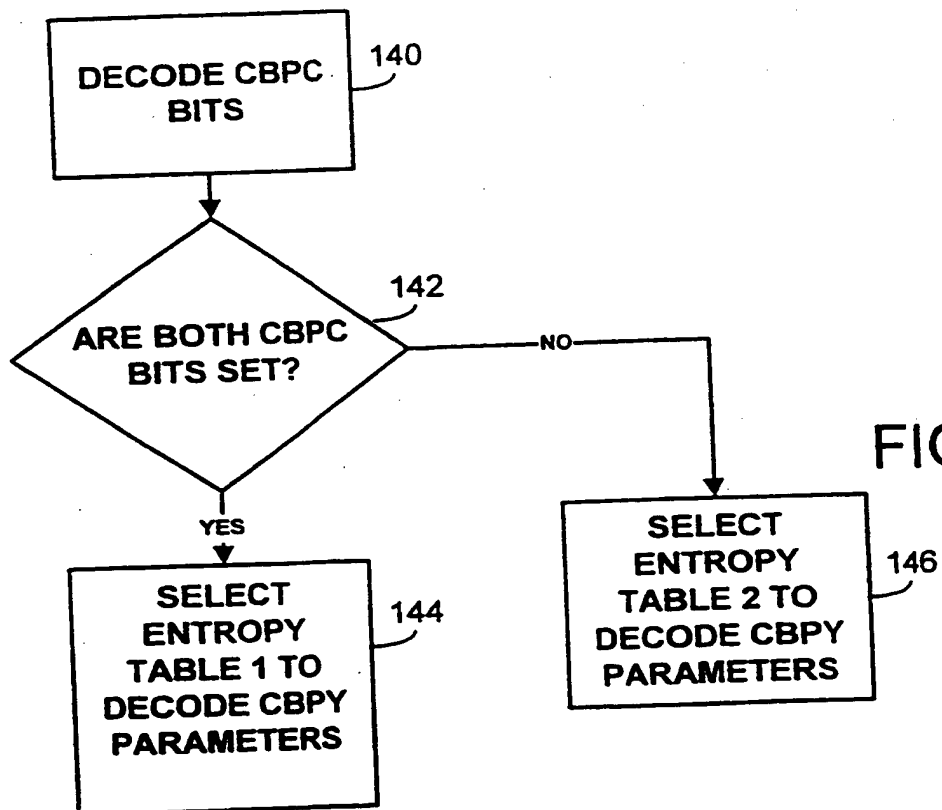
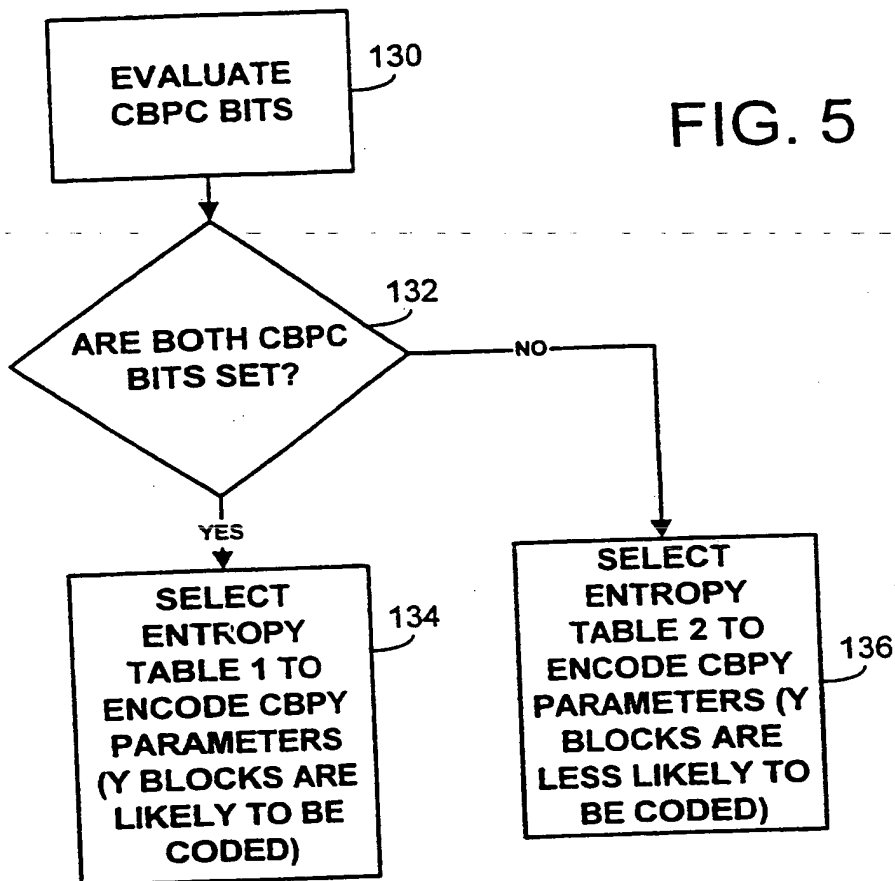


FIG. 6

INTERNATIONAL SEARCH REPORT

In tional Application No

PCT/US 98/20573

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04N7/36 H04N7/50 ---

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	AD HOC GROUP ON MPEG-4 VIDEO VM EDITING: "MPEG-4 VIDEO VERIFICATION MODEL VERSION 7.0 - ISO/IEC JTC1/SC29/WG11 MPEG97/N1642 - APPENDIX A : COMBINED MOTION SHAPE TEXTURE CODING - PAGES 195-203" April 1997, VIDEO GROUP (DRAFT IN PROGRESS), INTERNATIONAL STANDARDS ORGANIZATION, CODING OF MOVING PICTURES AND ASSOCIATED AUDIO INFORMATION XP002090865 see section 6.1.1 ---	1-20
A	EP 0 771 114 A (HYUNDAI ELECTRONICS IND) 2 May 1997 see page 3, line 3 - page 3, line 16 see claims 1-5 --- -/--	1-20

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

22 January 1999

Date of mailing of the international search report

03/02/1999

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Fassnacht, C

INTERNATIONAL SEARCH REPORT

Int lional Application No

PCT/US 98/20573

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 786 907 A (TEXAS INSTRUMENTS INC) 30 July 1997 see column 8, line 36 - column 9, line 4 ---	1-12,20
A	EP 0 707 425 A (OKI ELECTRIC IND CO LTD) 17 April 1996 see column 12, line 58 - column 13, line 45 ---	1-12,20
A	US 5 552 832 A (ASTLE BRIAN) 3 September 1996 see column 7, line 66 - column 8, line 44 see figures 4-6 ---	1-12,20
A	EP 0 588 653 A (FUJITSU LTD) 23 March 1994 see column 56, line 3 - column 57, line 8 see figures 6,10,33 see claim 44 ---	1-12,20
A	WEBB J L H: "POSTPROCESSING TO REDUCE BLOCKING ARTIFACTS FOR LOW BIT-RATE VIDEO CODING USING CHROMINANCE INFORMATION" 16 September 1996, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (IC, LAUSANNE, SEPT. 16 - 19, 1996, VOL. VOL. 2, PAGE(S) 9 - 12, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS XP000733222 see section "3. Block classification using chrominance information" -----	13-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/20573

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0771114 A	02-05-1997	CN 1159714 A JP 9186995 A NO 964495 A	17-09-1997 15-07-1997 28-04-1997
EP 0786907 A	30-07-1997	NONE	
EP 0707425 A	17-04-1996	JP 8111867 A US 5835144 A	30-04-1996 10-11-1998
US 5552832 A	03-09-1996	US 5557330 A	17-09-1996
EP 0588653 A	23-03-1994	JP 6153180 A	31-05-1994

THIS PAGE BLANK (USPTO)